

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
31 January 2002 (31.01.2002)

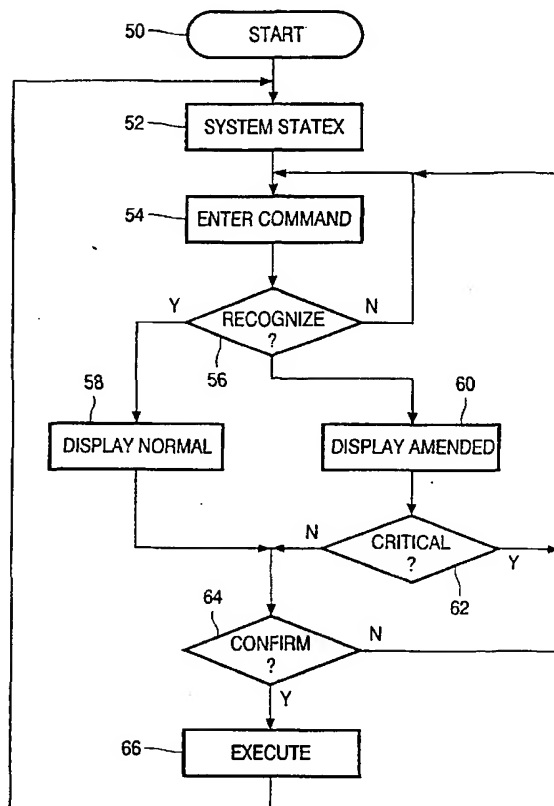
PCT

(10) International Publication Number
WO 02/09093 A1

- (51) International Patent Classification⁷: **G10L 15/22** (72) Inventors: **GEURTS, Lucas, J., F.**; Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL). **KAUFHOLZ, Paul, A., P.**; Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL).
- (21) International Application Number: **PCT/EP01/07847**
- (22) International Filing Date: **6 July 2001 (06.07.2001)** (74) Agent: **GROENENDAAL, Antonius, W., M.**; Internationaal Octrooibureau B.V., Prof Holstlaan 6, NL-5656 AA Eindhoven (NL).
- (25) Filing Language: **English**
- (26) Publication Language: **English** (81) Designated States (*national*): **CN, JP, KR.**
- (30) Priority Data: **00202607.8** **20 July 2000 (20.07.2000)** **EP** (84) Designated States (*regional*): **European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR).**
- (71) Applicant: **KONINKLIJKE PHILIPS ELECTRONICS N.V. [NL/NL]**; Groenewoudseweg 1, NL-5621 BA Eindhoven (NL). **Published:**
— *with international search report*

[Continued on next page]

(54) Title: **FEEDBACK OF RECOGNIZED COMMAND CONFIDENCE LEVEL**



(57) Abstract: An interactive user facility is operated through inputting voiced user commands, recognizing commands, executing recognized commands, and generating user feedback as regarding the progress of the operating. In particular, the recognizing asserts an associated confidence level and generates the user feedback through for a questionable command recognition presenting audio and/or video amending of the feedback with respect to both a correct recognition and with respect to a faulty recognition.

WO 02/09093 A1



For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

Feedback of recognized command confidence level

BACKGROUND OF THE INVENTION

The invention relates to a method as recited in the preamble of Claim 1. Voice control of interactive user facilities is being considered as an advantageous control mode in various environments, such as for handicapped persons, for machine operators using their hands for other tasks, as well as for the general public who find such feature an extremely advantageous convenience. However, speech recognition is not yet perfect. Recognition errors come in various categories: **deletion** errors will fail to recognize a speech item, **insertion** errors will recognize an item that has not effectively been uttered, and **substitution** errors will recognize another item than the one that has effectively been uttered. Especially, the last two situations may cause a faulty operation of the facility in question, and may therefore cause loss of information or money, incurred undue costs, malfunction of the facility, and possibly dangerous accidents. However, also deletion may cause nuisance. Feedback to the user can be presented by displaying the recognized phrase. The inventors have realized that the speech recognition is associated with various confidence levels, in that the recognition may be considered correct, questionable, or faulty, and that the overall user interaction would benefit from presenting an indication of the various levels representing such confidence, in association with executing the command or otherwise. Such feedback would indicate to a user person a particular speech item that should be repeated, possibly while being spoken with improved pronunciation or loudness, or rather, that the whole command needs improvement.

SUMMARY TO THE INVENTION

In consequence, amongst other things, it is an object of the present invention to improve the user interface of such an interactive user facility through representing various such confidence levels with respect to the recognizing of at least selected commands.

Now therefore, according to one of its aspects the invention is characterized according to the characterizing part of Claim 1.

The invention also relates to a device arranged for implementing a method as claimed in Claim 1. Further advantageous aspects of the invention are recited in dependent Claims.

5 BRIEF DESCRIPTION OF THE DRAWING

These and further aspects and advantages of the invention will be discussed more in detail hereinafter with reference to the disclosure of preferred embodiments, and in particular with reference to the appended Figures that show:

Figure 1, a general speech-enhanced user facility;

10 Figure 2, a flow chart illustrating a method embodiment of the present invention.

DETAILED DESCRIPTION OF PREFERRED EMBODIMENTS

Figure 1 illustrates a general speech-enhanced user facility for practicing the present invention. Block 20 represents the prime data processing module, such as a personal computer. Block 26 is a device for mechanical user input, such as keyboard, mouse, joystick or the like. Also shown are general block 22 for inputting data, such as memory or network, and general block 24 for outputting data, such as memory, network or printer. Block 34 represents an optional external facility that should be user-controlled, and which interfaces to the computer by I/O devices 36, such as sensors and actuators. The facility may be a consumer audio-video product, a factory automation facility, a motor vehicle information system or another data processing product. The latter external facility need not be present, inasmuch as user control by speech may be effected on the computer itself. Alternatively, the computer itself can form part of the external facility, for example an audio/video apparatus. Finally, there is a bi-directional audio interface with speech input 32 and speech or audio output 30. As will become evident, audio/speech output is optional.

Figure 2 represents a flow chart illustrating a method embodiment of the present invention. In start block 50 the data processing is activated, together with the assigning of the necessary facilities such as memory. In block 52 the system goes to a state indicated as "STATE X" that represents any applicable situation wherein the recognition of a user speech utterance is relevant for the operation. The attaining of this state so far is irrelevant for the present invention. Also, various further non-relevant aspects of the Figure have been suppressed, such as the eventual leaving of the flow chart. Now, in block 54 the user will enter a speech command, which the system then undertakes to recognize, which

recognizing can have an associated level of confidence. In block 56 the actual confidence level of the recognizing is assessed.

Now first, the recognition may be effectively correct, which will lead to displaying the recognized command in a normal manner, block 58. The system then asks the user to confirm, block 64. For this purpose, the system may allow a particular time span of a few seconds, so that non-confirming and not timely confirming will have the same effect. If validly confirmed, the command is executed, block 66, and the system reverts to block 52, that now represents the next system state "STATE X+1" wherein the recognition of a user speech utterance is relevant for the operation. If for a particular command no confirming is deemed necessary, the system would proceed immediately to block 66. For simplicity, the situation wherein **no** such speech input would be required in the applicable state has been ignored.

Second, the recognition may be faulty. This may be caused by various effects or circumstances. The speech itself may deficient, such as through being soft or inarticulate or occurring in a noisy environment. Also, the **content** of the speech may be deficient, such as through lacking a particular parameter value. Another problem is caused by superfluous speech elements (ahum!), wrong or inappropriate words or any other sort of lexical or semantic deficiencies. In these cases, the system goes back to block 54. This return may be associated by displaying what has been recognized if anything of the command in question, by a particular audio noise on item 30 in Figure 1 that indicates such return, by a particular expression in speech such as by displaying a request "repeat command", or by a textual display of the same. In certain situations, no return is executed, for example, through executing a default action.

Third, the recognition may have a questionable confidence level, which has been indicated by ?. This will cause an amended display of the recognized command in question with respect to the display effected in the case of correct recognition, block 60. The amending may pertain to the whole command, or only to the particular word or words of a plural-word command that effectively have a low confidence level. The amendment may be effected by another font or font size, a bold display versus normal, blinking, color, or any of various attention-grabbing mechanisms that by themselves have been common in text display. A particular feature would be the showing of an associated icon, such as an unsmiling face. Alternatively or in combination therewith, the system may produce an audio feedback that differs from the audio feedback in the case of reliable recognition in block 56, and also differs from the audio feedback in the case of faulty recognition in block 56. In

block 62 the system detects existence of a critical situation. This may pertain to an actual or expected command that by itself is critical, or in that the questionable recognition itself would bring about a critical situation. Executing a critical command could ensue high costs such as for example, by transferring money, or by starting a welding operation that cannot be
5 terminated halfway. Deleting of information may or may not be critical, as the case be. If critical however, the system reverts to block 54 for a new speech command entry. If non-critical, the system asks for confirm in block 64, and the situation corresponds to correct recognition. In certain situations, the questionable recognition would need just signaling thereof to a user person, as an urge to improve the quality of the voice commands, such as by
10 better pronunciation.

The procedure may be amended in various manners. The confidence may have more than three levels, each with their associated display amending, categorizing of which is critical and which is not, partial or full repeating of an uttered command, and the like. Persons skilled in the art will appreciate various amendments to the preferred embodiment
15 disclosed supra that would bring about the advantages of the invention, without departing from its scope as defined by the appended Claims hereinafter.

CLAIMS:

1. A method for operating an interactive user facility through inputting voiced user commands, recognizing such commands, executing such recognized commands, and generating user feedback as regarding the progress of such operating,
said method being characterized by in such recognizing asserting an associated
5 confidence level and generating such user feedback through for a questionable command recognition presenting audio and/or video amending of such feedback both with respect to a correct recognition and with respect to a faulty recognition.
2. A method as claimed in Claim 1, wherein such presenting is based on selective
10 amending of a textual display of a recognized command with respect to a standard display.
3. A method as claimed in Claim 1, wherein such presenting is based on selective amending of an audio feedback item with respect to a standard audio feedback.
- 15 4. A method as claimed in Claim 1, wherein such presenting is based on selective iconizing with respect to a standard display.
5. A method as claimed in Claim 1, wherein a questionable recognition stalls execution of at least certain of such recognized commands.
20
6. An apparatus being arranged for practicing a method as claimed in Claim 1 for operating an interactive user facility and having input means for receiving voiced user commands, recognizing means for recognizing such commands, execution means for executing such recognized commands, and feedback generating means for generating user
25 feedback as regarding the progress of such operating,
said apparatus being characterized by having asserting means for in such recognizing asserting an associated confidence level and feeding said feedback generating means for generating such user feedback for a questionable command recognition through

presenting audio and/or video amending of such feedback both with respect to a correct recognition and with respect to a faulty recognition.

7. An apparatus as claimed in Claim 6, and having amending means for
5 selectively amending a textual display of a recognized command with respect to a standard display.
8. An apparatus as claimed in Claim 6, and having amending means for
selectively amending an audio feedback item with respect to a standard audio feedback.
- 10 9. An apparatus as claimed in Claim 6, and having amending means for selective
iconizing with respect to a standard display.
10. An apparatus as claimed in Claim 6, and having stall means activated by a
15 questionable recognition for stalling execution of at least certain of such recognized
commands.

1/2

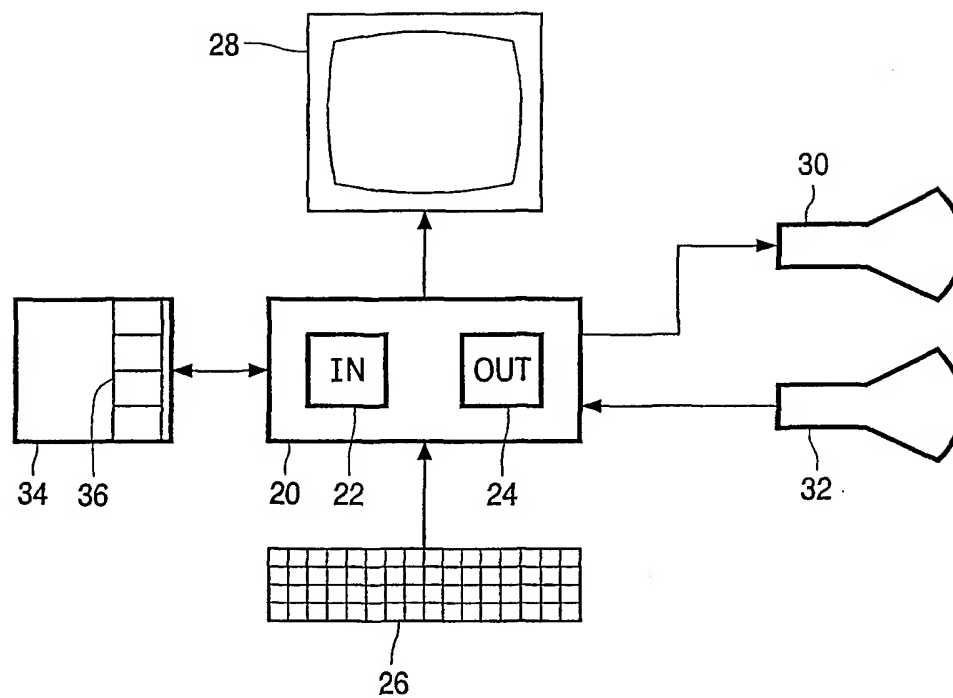


FIG. 1

2/2

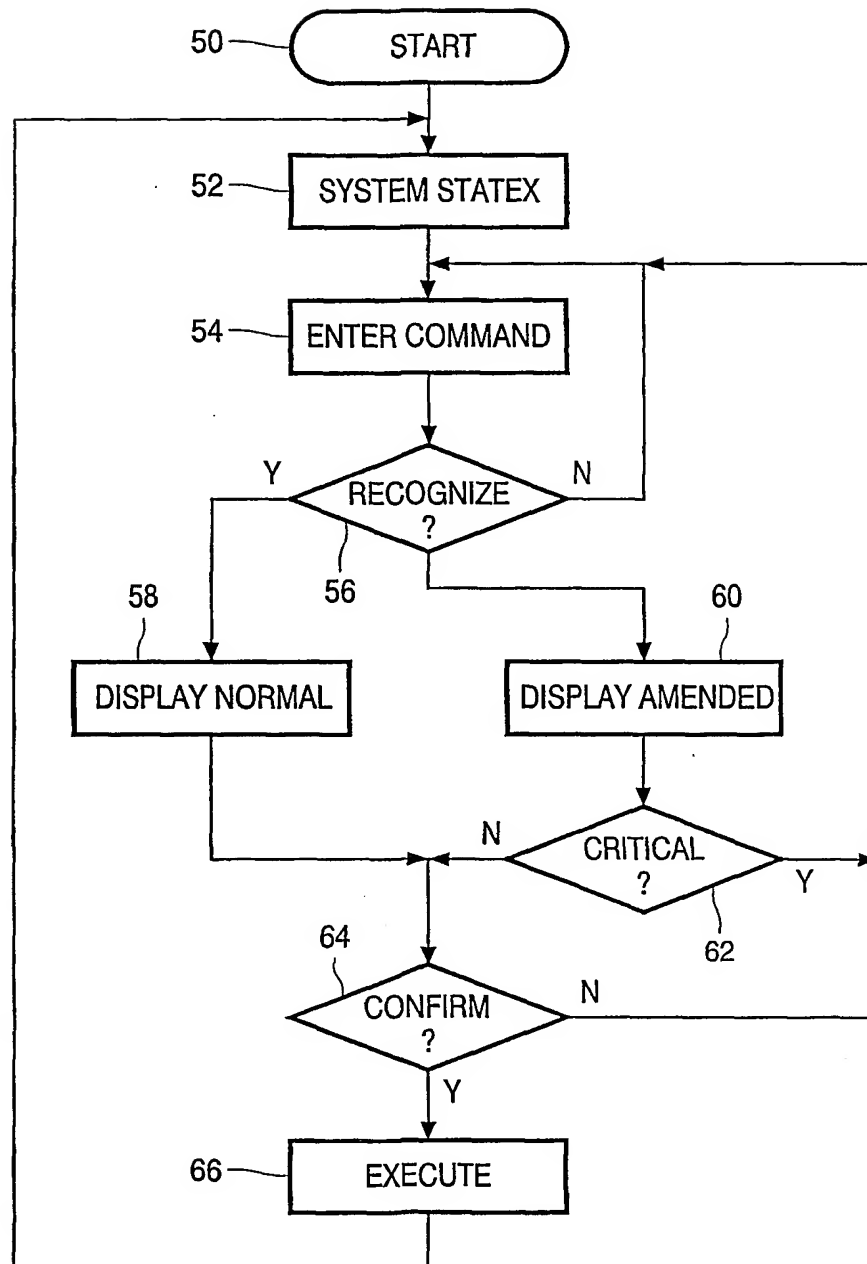


FIG. 2

INTERNATIONAL SEARCH REPORT

In International Application No
PCT/EP 01/07847A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 G10L15/22

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 G10L G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, PAJ, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EP 0 651 372 A (AT. & T CORP) 3 May 1995 (1995-05-03) figure 5 column 7, line 34 -column 8, line 13	1,3,5,6, 8,10
Y	----	4,9
X	EP 0 924 687 A (IBM) 23 June 1999 (1999-06-23) column 3, line 3 - line 22 column 4, line 24 - line 34	1,2,6,7
X	EP 0 957 470 A (PHILIPS PATENTVERWALTUNG ;KONINKL PHILIPS ELECTRONICS NV (NL)) 17 November 1999 (1999-11-17) abstract ----- -/--	1,2,6,7

☒ Further documents are listed in the continuation of box C.☒ Patent family members are listed in annex.

* Special categories of cited documents :

A document defining the general state of the art which is not considered to be of particular relevance

E earlier document but published on or after the international filing date

L document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

O document referring to an oral disclosure, use, exhibition or other means

P document published prior to the international filing date but later than the priority date claimed

T later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

X document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

Y document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

Z document member of the same patent family

Date of the actual completion of the international search

25 September 2001

Date of mailing of the international search report

08/10/2001

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Ramos Sánchez, U

INTERNATIONAL SEARCH REPORT

II International Application No
PCT/EP 01/07847

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 5 864 815 A (BELGAU RONALD A ET AL) 26 January 1999 (1999-01-26) figures 9,10 column 5, line 36 -column 6, line 4 ----	4,9
A	RHYNE J R ET AL: "RECOGNITION-BASED USER INTERFACES" ADVANCES IN HUMAN COMPUTER INTERACTION, XX, XX, no. 4, 1993, pages 191-250, XP002129803 page 216, last paragraph -page 217, paragraph 1 -----	1-10
A	EP 0 850 673 A (SEGA ENTERPRISES KK) 1 July 1998 (1998-07-01) page 7, line 18 - line 23 -----	4,9

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/EP 01/07847

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
EP 0651372	A	03-05-1995	US 5566272 A CA 2131600 A1 CN 1115902 A EP 0651372 A2 JP 7181994 A	15-10-1996 28-04-1995 31-01-1996 03-05-1995 21-07-1995
EP 0924687	A	23-06-1999	US 6006183 A EP 0924687 A2 JP 11231891 A	21-12-1999 23-06-1999 27-08-1999
EP 0957470	A	17-11-1999	DE 19821422 A1 CN 1238489 A EP 0957470 A2 JP 11352992 A	18-11-1999 15-12-1999 17-11-1999 24-12-1999
US 5864815	A	26-01-1999	NONE	
EP 0850673	A	01-07-1998	EP 0850673 A1 CN 1197404 A WO 9802223 A1	01-07-1998 28-10-1998 22-01-1998